

10/554387
JC09 Rec'd PCT/PTO 25 OCT 2003

1
SEQUENCE LISTING

<110> Shaaltiel, Yoseph
Baum, Gideon
Sharon Hashmueli
Ayala Lewkowicz
Bartfeld, Daniel

<120> PRODUCTION OF HIGH MANNOSE PROTEINS IN PLANT CULTURE

<130> 30570

<160> 14

<170> PatentIn version 3.3

<210> 1
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> ER signal peptide

<400> 1

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
1 5 10 15

Leu Ser Ser Ala Glu Phe
20

<210> 2
<211> 7
<212> PRT
<213> Artificial sequence

<220>
<223> Vacuolar targeting signal from Tobacco chitinase A

<400> 2

Asp Leu Leu Val Asp Thr Met
1 5

<210> 3
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Single strand DNA oligonucleotide

<400> 3
cagaattcgc ccggccctgc a

21

<210> 4
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Single strand DNA oligonucleotide

<400> 4
ctcagatctt ggcgatgcca ca

22

<210> 5
<211> 19
<212> DNA

<213> Artificial sequence		
<220>		
<223> Single strand DNA oligonucleotide		
<400> 5		19
ctcagaagac cagaggct		
<210> 6		
<211> 17		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Single strand DNA oligonucleotide		
<400> 6		17
caaagcggcc atcgtgc		
<210> 7		
<211> 1491		
<212> DNA		
<213> Homo sapiens		
<400> 7		
gccccccctt gcatccctaa aagcttcggc tacagctcgg tgggtgtgt ctgcaatgcc	60	
acatactgtg actcctttga ccccccggacc tttcctgccc ttggcacctt cagccgctat	120	
gagagtacac gcagtggcg acggatggag ctgagtatgg ggcccatcca ggctaatcac	180	
acgggcacag gcctgtact gaccctgcag ccagaacaga agttccagaa agtgaaggaa	240	
tttgggggggccatgacaga tgctgctgct ctcaacatcc ttgcctgtc accccctgccc	300	
caaaatttgc tacttaaattc gtacttctct gaagaaggaa tcggatataa catcatccgg	360	
gtacccatgg ccagctgtga ctctccatc cgcacctaca cctatgcaga caccctgtat	420	
gatttccagt tgcacaactt cagcctccca gaggaagata ccaagctcaa gataccctg	480	
attcaccgag ccctgcagtt ggcccagcgt cccgtttcac tccttgcag cccctggaca	540	
tcacccactt ggctcaagac caatggagcg gtgaatggga aggggtcact caagggacag	600	
cccgagaca tctaccacca gacctggggcc agatactttg tgaagttctt ggatgcctat	660	
gctgagcaca agttacagtt ctgggcagtg acagctgaaa atgagcccttc tgctggcgt	720	
ttgagtggat acccccttcca gtgcctggc ttcacccctg aacatcagcg agacttcatt	780	
gcccgtgacc taggtcctac cctcgccaaac agtactcacc acaatgtccg cctactcatg	840	
ctggatgacc aacgcttgct gctccccac tggccaaagg tggtaactgac agacccagaa	900	
gcagctaaat atgttcatgg cattgctgta cattggtacc tggactttctt ggctccagcc	960	
aaagccaccc tagggagac acaccgcctg ttccccaaaca ccatgctttt tgcctcagag	1020	
gcctgtgtgg gctccaaagtt ctgggagcag agtgtgcggc taggctcctg ggatcgaggg	1080	
atgcagttaca gccacagcat catcacgaac ctccgttacc atgtggtcgg ctggaccgac	1140	
tggAACCTTG CCCTGAACCC CGAAGGAGGA CCCAATTGGG TGCCTAATT TGTGACAGT	1200	
CCCATCATG TAGACATCAC CAAGGACACG TTTACAAAC AGCCCATGTT CTACCACCTT	1260	
GGCCACTTCA GCAAGTTCAT TCCTGAGGGC TCCCAGAGAG TGGGGCTGGT TGCCAGTCAG	1320	
AAGAACGACC TGGACGAGT GGCACTGATG CATCCCGATG GCTCTGCTGT TGTGGTCGTG	1380	

ctaaaccgct cctctaagga tgtgcctt accatcaagg atcctgctgt gggcttcctg 1440
 gagacaatct cacctggcta ctccattcac acctacctgt ggcacatgcca g 1491

<210> 8
 <211> 335
 <212> PRT
 <213> Homo sapiens

<400> 8

Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly Tyr Ser Ser Val Val Cys
 1 5 10 15

Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe Asp Pro Pro Thr Phe Pro
 20 25 30

Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser Thr Arg Ser Gly Arg Arg
 35 40 45

Met Glu Leu Ser Met Gly Pro Ile Gln Ala Asn His Thr Gly Thr Gly
 50 55 60

Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys Phe Gln Lys Val Lys Gly
 65 70 75 80

Phe Gly Gly Ala Met Thr Asp Ala Ala Leu Asn Ile Leu Ala Leu
 85 90 95

Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys Ser Tyr Phe Ser Glu Glu
 100 105 110

Gly Val Arg Leu Leu Met Leu Asn Asp Gln Arg Leu Leu Leu Pro His
 115 120 125

Trp Ala Lys Val Val Leu Thr Asp Pro Glu Ala Ala Lys Tyr Val His
 130 135 140

Gly Ile Ala Val His Trp Tyr Leu Asp Phe Leu Ala Pro Ala Lys Ala
 145 150 155 160

Thr Leu Gly Glu Thr His Arg Leu Phe Pro Asn Thr Met Leu Phe Ala
 165 170 175

Ser Glu Ala Cys Val Gly Ser Lys Phe Trp Glu Gln Ser Val Arg Leu
 180 185 190

Gly Ser Trp Asp Arg Gly Met Gln Tyr Ser His Ser Ile Ile Thr Asn
 195 200 205

Leu Leu Tyr His Val Val Gly Trp Thr Asp Trp Asn Leu Ala Leu Asn
 210 215 220

Pro Glu Gly Gly Pro Asn Trp Val Arg Asn Phe Val Asp Ser Pro Ile
 225 230 235 240

Ile Val Asp Ile Thr Lys Asp Thr Phe Tyr Lys Gln Pro Met Phe Tyr

245 250

255

His Leu Gly His Phe Ser Lys Phe Ile Pro Glu Gly Ser Gln Arg Val
 260 265 270

Gly Leu Val Ala Ser Gln Lys Asn Asp Leu Asp Ala Val Ala Leu Met
 275 280 285

His Pro Asp Gly Ser Ala Val Val Val Val Leu Asn Arg Ser Ser Lys
 290 295 300

Asp Val Pro Leu Thr Ile Lys Asp Pro Ala Val Gly Phe Leu Glu Thr
 305 310 315 320

Ile Ser Pro Gly Tyr Ser Ile His Thr Tyr Leu Trp His Arg Gln
 325 330 335

<210> 9
 <211> 338
 <212> DNA
 <213> Cauliflower mosaic virus

<400> 9
 ttttcacaaa gggtaatatac gggaaacctc ctcggattcc attgcccagc tatctgtcac 60
 ttcatcgaaa ggacagttaga aaaggaaggt ggctcctaca aatgccatca ttgcgataaa 120
 ggaaaggctta tcgttcaaga tgcctctacc gacagtggtc ccaaagatgg acccccaccc 180
 acgaggaaca tcgtggaaaa agaagacgtt ccaaccacgt cttcaaagca agtggattga 240
 tgtgatatatct ccactgacgt aaggatgac gcacaatccc actatccttc gcaagaccct 300
 tcctctatat aaggaagttc atttcatttg gagaggac 338

<210> 10
 <211> 66
 <212> DNA
 <213> Artificial sequence
 <220>
 <223> Nucleic acid sequence encoding the ER signal peptide
 <400> 10
 atgaagacta atcttttct ctttctatc ttttacttc tcctatcatt atcctcgccc 60
 gaattc 66

<210> 11
 <211> 21
 <212> DNA
 <213> Artificial sequence
 <220>
 <223> Nucleic acid sequence encoding the vacuolar targeting sequence
 <400> 11
 gatcttttag tcgatactat g 21

<210> 12
 <211> 167
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Nucleic acid sequence of the Agrobacterium tumefaciens terminator

<220>
 <221> misc_feature
 <222> (162)..(162)
 <223> n is a, c, g, or t

<400> 12
 taatttcatg atctgtttg ttgtattccc ttgcaatgca gggcctaggg ctatgaataa 60
 agttaatgtg tgaatgtgt aatgtgtat tgtgacctga agggatcacg actataatcg 120
 tttataataa acaaagactt tgtcccaaaa accccccccc cnngcaga 167

<210> 13
 <211> 2186
 <212> DNA
 <213> Artificial sequence

<220>
 <223> nucleic acid sequence encoding high mannose human
 glucocerebrosidase (GCD)

<220>
 <221> misc_feature
 <222> (2181)..(2181)
 <223> n is a, c, g, or t

<400> 13
 ttttcacaaa gggtaatatac gggaaacctc ctcggattcc attgcccagc tatctgtcac 60
 ttcatcgaaa ggacagtaga aaaggaaggt ggctcctaca aatgcccata ttgcgataaa 120
 ggaaaggcta tcgttcaaga tgcctctacc gacagtggtc ccaaagatgg acccccaccc 180
 acgaggaaca tcgtggaaaa agaagacgtt ccaaccacgt cttcaaaagca agtggattga 240
 tgtgatatct ccactgacgt aaggatgac gcacaatccc actatccttc gcaagaccct 300
 tcctctatat aaggaagttc atttcatttg gagaggacag gcttcttgag atccttcaac 360
 aattaccaac aacaacaaac aacaacaaac attacaatta ctatccataa ttacagtca 420
 gggatccaag gagatataac aatgaagact aatcttttc tctttctcat ctttcactt 480
 ctccatcatat ttccttcggc cgaattcgcc cggccctgca tccctaaag cttcggtac 540
 agctcggtgg tgtgtgtctg caatgccaca tactgtact cctttgaccc cccgaccctt 600
 cctgccttg gtaccttcag ccgctatgag agtacacgca gtggcgacg gatggagctg 660
 agtatggggc ccatccaggc taatcacacg ggcacaggcc tgctactgac cctgcagcca 720
 gaacagaagt tccagaaagt gaaggatgg ggagggccca tgacagatgc tgctgcttc 780
 aacatccttg ccctgtcacc ccctgccc aaatttgcata ttaaatcgta cttctctgaa 840
 gaaggaatcg gatataacat catccggta cccatggcc gctgtgactt ctccatccgc 900
 acctacacct atgcagacac ccctgatgat ttccagttgc acaacttcag cctcccagag 960
 gaagatacca agctcaagat acccctgatt caccgagccc tgcagttggc ccagcgtccc 1020
 gtttcactcc ttgccagccc ctggacatca cccacttggc tcaagaccaa tggagcggtg 1080
 aatgggaagg ggtcaactaa gggacagccc ggagacatctt accaccagac ctggggccaga 1140
 tactttgtga agttccttggaa tgcctatgct gaggcacaagt tacagttctg ggcagtgaca 1200

gctgaaaatg agccttctgc tggctgttg agtggatacc ccttccagtg cctgggcttc	1260
acccctgaac atcagcgaga cttcattgcc cgtgacctag gtcctaccct cgccaaacagt	1320
actcaccaca atgtccgcct actcatgctg gatgaccaac gcttgctgct gccccactgg	1380
gcaaagggtgg tactgacaga cccagaagca gctaaatatg ttcatggcat tgctgtacat	1440
tggtacctgg actttctggc tccagccaaa gccaccctag gggagacaca cgcctgttc	1500
cccaacacca tgctcttgc ctcaagggcc tgggtggct ccaagttctg ggagcagagt	1560
gtgcggctag gctccctgggaa tcgagggatg cagtacagcc acagcatcat cacgaacctc	1620
ctgtaccatg tggtcggctg gaccgactgg aaccttgcctc tgaaccccgaa aggaggaccc	1680
aattgggtgc gtaactttgt cgacagtccc atcattgttag acatcaccaa ggacacgttt	1740
tacaaacacg ccatgttcta ccacccctggc cacttcagca agttcattcc tgagggctcc	1800
cagagagtgg ggctgggtgc cagtcagaag aacgacactgg acgcagtggc actgatgcat	1860
cccgatggct ctgctgtgtt ggtcgtgcta aaccgctcct ctaaggatgt gccttacc	1920
atcaaggatc ctgctgtggg ctccctggag acaatctcac ctggctactc cattcacacc	1980
tacctgtggc atcgccaaga tcttttagtc gatactatgt aatttcatga tctgtttgt	2040
tgtattccct tgcaatgcag ggcctaggc tatgaataaa gttaatgtgt gaatgtgtga	2100
atgtgtgatt gtgacctgaa gggatcacga ctataatcg ttataataaa caaagacttt	2160
gtccccaaaaa cccccccccc ngcaga	2186

<210> 14
 <211> 526
 <212> PRT
 <213> Artificial sequence

<220>
 <223> High mannose human glucocerebrosidase (GCD)

<400> 14

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser			
1	5	10	15

Leu Ser Ser Ala Glu Phe Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly		
20	25	30

Tyr Ser Ser Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe		
35	40	45

Asp Pro Pro Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser		
50	55	60

Thr Arg Ser Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln Ala			
65	70	75	80

Asn His Thr Gly Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys		
85	90	95

Phe Gln Lys Val Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala Ala		
100	105	110

Leu Asn Ile Leu Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys
 115 120 125

Ser Tyr Phe Ser Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val Pro
 130 135 140

Met Ala Ser Cys Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp Thr
 145 150 155 160

Pro Asp Asp Phe Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp Thr
 165 170 175

Lys Leu Lys Ile Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln Arg
 180 185 190

Pro Val Ser Leu Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu Lys
 195 200 205

Thr Asn Gly Ala Val Asn Gly Lys Gly Ser Leu Lys Gly Gln Pro Gly
 210 215 220

Asp Ile Tyr His Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu Asp
 225 230 235 240

Ala Tyr Ala Glu His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu Asn
 245 250 255

Glu Pro Ser Ala Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu Gly
 260 265 270

Phe Thr Pro Glu His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly Pro
 275 280 285

Thr Leu Ala Asn Ser Thr His His Asn Val Arg Leu Leu Met Leu Asp
 290 295 300

Asp Gln Arg Leu Leu Leu Pro His Trp Ala Lys Val Val Leu Thr Asp
 305 310 315 320

Pro Glu Ala Ala Lys Tyr Val His Gly Ile Ala Val His Trp Tyr Leu
 325 330 335

Asp Phe Leu Ala Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg Leu
 340 345 350

Phe Pro Asn Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser Lys
 355 360 365

Phe Trp Glu Gln Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met Gln
 370 375 380

Tyr Ser His Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly Trp
 385 390 395 400

Thr Asp Trp Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp Val
 405 410 415

Arg Asn Phe Val Asp Ser Pro Ile Ile Val Asp Ile Thr Lys Asp Thr
420 425 430

Phe Tyr Lys Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys Phe
435 440 445

Ile Pro Glu Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys Asn
450 455 460

Asp Leu Asp Ala Val Ala Leu Met His Pro Asp Gly Ser Ala Val Val
465 470 475 480

Val Val Leu Asn Arg Ser Ser Lys Asp Val Pro Leu Thr Ile Lys Asp
485 490 495

Pro Ala Val Gly Phe Leu Glu Thr Ile Ser Pro Gly Tyr Ser Ile His
500 505 510

Thr Tyr Leu Trp His Arg Gln Asp Leu Leu Val Asp Thr Met
515 520 525